DOCUMENT RESUME

ED 414 170 SE 060 860

AUTHOR Miller-Whitehead, Marie
TITLE A Longitudinal Analysis of Science Scale Scores Grades 2-8 in Tennessee for 1992-1996.
PUB DATE 1997-10-00
NOTE 25p.; For a related study, see SE 060 722.
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *Academic Standards; Analysis of Variance; *Educational Assessment; Elementary Education; Longitudinal Studies; *Norm Referenced Tests; *Science Education; *Scores
IDENTIFIERS Tennessee

ABSTRACT

In light of the continuing debate over the relative merits of various ways of assessing student achievement, it seems appropriate to examine longitudinal data which reflect student performance on one measure as a basis for recognizing and utilizing the measure's particular strengths. This study is the follow-up to an earlier study conducted using data from the state of Tennessee for student scale scores in science for the years 1990-1994. The current data set consisted of school system level science scale scores on the CTBS/4 science test, grades 2 through 8, for each of the 138 Tennessee school systems for the years 1992-1996. The normed portion of the science subtest of the CTBS/4 consists of 20 items with four possible answers for each item. Examination of the descriptives revealed an increase in the mean of science scale scores for grades 2-8 each year except 1993, which showed a "negative gain." This population represents the remaining members of the cohort of students (grades 4-8) included in the mean score for 1991, which also had a "negative gain." This finding seems to indicate that teacher effect on student achievement may be both cumulative and residual. Contains 21 references. (Author/PVD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 414 170

A longitudinal analysis of science scale scores grades 2 - 8 in Tennessee for 1992 - 1996

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

M. Miller-Whitehead

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Marie Miller-Whitehead
Education Consultant
October 1997

SE 060 810

BEST COPY AVAILABLE

In light of the continuing debate over the relative merits of various ways of assessing student achievement, including criterion referenced tests, performance-based assessments, norm-referenced tests, and portfolio evaluation to mention only a few, it seems appropriate to examine more closely longitudinal data which reflect student performance on one measure as a basis for recognizing and utilizing its particular strengths. For example, scoring of performance assessments for students in K-12 continues to evolve and become more sophisticated as more and more teachers gain expertise in its usage; however, there are certain advantages to be gained from an examination of data which have been recorded on a consistent measurement over a long period of time. Those states, school systems, or institutions which are fortunate enough to have in place mechanisms for collecting and disseminating such data can provide invaluable information for school administrators, policymakers, and the community. The value of this information lies not only in what is answered by the data but in what remains to be answered. As accountability mechanisms become more pervasive and in some cases more closely tied to funding it is particularly critical to provide decisionmakers at every level with published findings and reports pertinent to policy and long range planning. For example, much of the data collected via Tennessee's accountability mechanism has provided researchers with empirical evidence on smaller class size (Achilles, Zaharias, & Nye, 1995; Finn & Achilles , 1990; Nye et al., 1992; Underwood &

Lumsden., 1994), multiage grouping (Nye, 1993), and teacher effects (Sanders & Rivers, 1996). Such initiatives as the National Assessment of Educational Progress (NAEP) have also provided policymakers with nationwide and state-by-state trends in student achievement accompanied by a variety of demographic and ethnic disaggregated data (Johnson et al., 1992; Campbell et al., 1996; Ballator, 1996; Bruschi & Anderson,1994). While opponents of accountability via standardized testing have been very vocal, there is no denying the fact that disaggregated data provided by these analyses have helped school districts to assess to what extent each subpopulation is achieving at the same level. Until other reliable assessments can be developed (be they performance or portfolio related) the data provided for policymakers, despite admitted drawbacks, are continuing to contribute to more effective schools and educational equity.

The focus of this analysis on student scores on the norm-referenced CTBS/4 test should not be taken to minimize the merits of any of the other measures of student performance, progress, and achievement. There are many resources available for those interested in a more indepth discussion of student assessment (McLean & Lockwood, 1996; Noble & Sawyer , 1992; O'Sullivan, 1995; Yepes Baraya, 1995), particularly in issues concerning reliability and validity of various instruments.

In addition to systems which are participating in NAEP projects and testing,

Tennessee has a state-mandated accountability system which provides for student testing each year in grades 2 - 8 in the five subject areas of reading, language arts, math, science, and social studies. There have been several detailed reports published on the Tennessee Comprehensive Assessment Program (Baker & Xu, 1995; Bock, Wolfe, and Fisher, 1996) which will provide those interested in comparison of various state accountability mechanisms with an overview of the program.

The present study is a followup to an earlier study conducted using state of Tennessee data for student scale scores in science for the years 1990-1994 (Miller-Whitehead, 1997). The findings of the earlier study indicated that mean science scale scores over grades 2 - 8 across the state of Tennessee had improved each year except 1991 (mean 721.42) from 1990 to 1994. However, there had also been an indication that while minimum mean scores had risen, maximum mean scores had declined. While this data is not necessarily indicative of a widespread plunge in the achievement of students on the higher end of the performance scale it does raise questions which each school system should be prepared to answer relative to its own student data.

The data set consisted of school system level science scale scores on the CTBS/4 science test, grades two through eight, for each of the 138 Tennessee school systems for the years 1992 to 1996. The normed portion of the science

subtest of the CTBS/4 consists of 20 itemswith four possible answers for each item. For Tennessee CTB computes IRT scale scores with a possible continuous range of values over grades K - 8 of from 0 - 999 (Tables 2, 3). The CTBS/4 technical manual provides additional information regarding benchmarks and test-retest reliability coefficients for each level of the science subtest. SEM for the IRT scores for each level of the test and information in respect to grade equivalencies for levels of the test is available in the Spring Norms manual. Those interested in specifics of the CTBS/4 may refer to published reviews (Bock et al., 1996; Miller, 1992; Hopkins, 1992) and the technical manual (CTB, undated). Five of the 138 systems were omitted due to grade configurations specific to individual systems; therefore the main analysis was conducted with data from the 133 systems which include grades two through eight. However, mean scores by year and mean scores by grade are mean scores for all systems which reported scores for the grades included in the analysis. The minimum and maximum mean scores reported also may reflect relatively small outlier systems or new specialized schools within systems with disproportionately large numbers of either high or low achieving students. The analysis was conducted using the SPSS for Windows 7.5 statistical software package.

An examination of the descriptives revealed an increase in the mean of science scale scores for grades 2-8 each year except 1993, which showed a "negative gain" (Table 1). This population represents the remaining members of the

cohort of students (grades 4 - 8) included in the mean score for year 1991 which also had a "negative gain." This finding would seem to reinforce that of Sanders and Rivers (1996) which determined that the teacher effect on student achievement may be both cumulative and residual. A comparison of the aggregate mean scores by grade level over the years 1990-1994 with the mean of scores over the years 1992-1996 showed an increase in mean for each grade level in the 1992-1996 scale scores (Table 2). The implication is that student achievement as measured by the CTBS/4 science test has improved and that for each grade level tested, students in later years are doing better on the average than their predecessors in public schools across the state of Tennessee. While this finding does not address how Tennessee science students compare with students across the U.S., these results are encouraging and point to the success of Tennessee's efforts to assure that all students receive a fair and equitable education.

To determine the statistical significance of these findings a within subjects MANOVA procedure was conducted with 5 levels for year and 7 levels for grade thus creating 35 new variables for mean science scale scores. The null hypothesis of the investigation was that there is no statistically significant difference in mean science scale scores across years or grade levels.

A preliminary examination of univariate parameters and 95% confidence intervals for the variables was conducted. The results of this analysis indicated that

the sample exceeded the norm for Year 93 and Year 94. The 95% confidence interval included 0 for both Year 95 and Year 96. For the Grade variable, Grade 3 exceeded the sample norm, Grade 4 was lower than the sample norm, Grade 6 had a 95% confidence which included 0, the Grade 7 sample was less than the norm, and the Grade 8 sample was more than the norm. An examination of univariate *F* tests for the variables showed that there were significant univariate tests for Year 93 and Year 94 and for Grade 3, Grade 4, Grade 5, Grade 7, and Grade 8. Year 94 and Year 95 had nonsignificant univariate *F*, as did Grade 6. These findings were in accordance with the examination of the univariate 95% confidence intervals. For the Year by Grade effect, univariate *F* tests showed nonsignificant univariate *F* tests for Year 93 Grade 3, for Year 93 Grade 8, for Year 94 Grade 3, for Year 94 Grade 8, for Year 95 Grade 3, and for Year 96 Grade 6. An examination of univariate 95% confidence intervals showed values exceeding the norm for Year 93 Grade 4, Year 93 Grade 7, Year 94 Grade 4, Year 94 Grade 5, Year 95 Grade 5, Year 95 Grade 6, Year 95 Grade 7, Year 95 Grade 8, Year 96 Grade 4, and Year 96 Grade 5. The sample values were poorer than the norm for Year 93 Grade 5, Year 93 Grade 6, Year 94 Grade 6, Year 94 Grade 7, Year 95 Grade 4, Year 96 Grade 3, Year 96 Grade 7, and Year 96 Grade 8. The 95% confidence intervals included 0 for Year 93 Grade 3, Year 93 Grade 8, Year 94 Grade 3, Year 94 Grade 8, Year 95 Grade 3, and Year 96 Grade 6. However, the *F* tests are not adjusted for number of variables in

the analysis. A stepdown analysis indicated that after controlling for Type I error there were significant effects for Year 93 Grade 4, Year 93 Grade 5, Year 93 Grade 6, Year 93 Grade 7, Year 93 Grade 8, Year 94 Grade 7, Year 95 Grade 6, Year 95 Grade 7, Year 95 Grade 8, Year 96 Grade 3, Year 96 Grade 5, and Year 96 Grade 8. Not surprisingly, the strength of association effect was greatest for the Grade variable (η^2 = .98) with the interaction of Year by Grade having practical significance at η^2 = .23. Practical significance of the effect of Year was marginal with an η^2 = .18.

Table 1

Mean science scale scores for grades 2 - 8 by year

1992		1993		1994		1995		1996	
M	N	*M*	N	*M*	N	*M*	N	*M*	N
723 90	956	723 14	956	724 34	956	726 47	957	728 90	958

Table 2

Mean science scale scores for Tennessee 1992 - 1996 by grade level

Grade	2	3	4	5	6	7	8
1990-1994	667.51	690.96	713.44	728.49	739.64	754.89	766.82
1992-1996	668.56 N=690	692.55 N=690	716.64 N=690	729.57 N=690	742.37 N=687	759.03 N=669	771.38 N=667

Table 3

Science scale score descriptives for Tennessee by grade level and by year

	N	*M*	min	max	variation	SD
SS92.2	138	667.01	630.70	697.50	155.65	12.48
SS92.3	138	690.57	662.90	720.30	118.94	10.91
SS92.4	138	718.57	695.90	739.50	60.55	7.78
SS92.5	138	727.22	690.90	774.20	81.25	9.01
SS92.6	138	734.00	699.00	763.90	106.63	10.33
SS92.7	138	757.62	730.10	781.30	66.33	8.14
SS92.8	138	768.07	740.80	795.30	90.01	9.49
SS93.2	138	662.57	627.90	692.90	157.98	12.57
SS93.3	138	686.48	653.50	717.40	119.89	10.95
SS93.4	138	716.46	681.60	741.40	119.28	10.92
SS93.5	138	726.97	699.60	751.20	72.34	8.51
SS93.6	138	746.42	705.70	775.60	106.53	10.32
SS93.7	138	754.55	729.70	779.00	61.99	7.87
SS93.8	138	770.67	747.80	794.60	54.75	7.40
SS94.2	138	674.56	625.20	714.60	166.74	12.91

Science scale scores Tennessee 1992 - 1996

	N	M	min	max	variation	SD
SS94 3	138	698.61	650.10	732.50	162.42	12.75
SS94 4	138	715.85	682.30	743.60	95.15	9.76
SS94 5	138	733.48	698.90	754.30	87.22	9.34
SS94 6	137	734.98	698.20	756.50	79.74	8.93
SS94 7	134	753.05	720.80	784.30	73.07	8.55
SS94 8	133	765.08	745.60	787.40	60.60	7.79
SS95 2	138	668.99	631.70	702.40	166.09	12.89
SS95 3	138	691.48	644.20	728.10	139.28	11.80
SS95 4	138	715.38	671.40	743.50	112.09	10.59
SS95 5	138	727.88	696.70	771.60	103.30	10.16
SS95 6	137	747.45	722.50	784.40	121.15	11.01
SS95 7	134	764.37	732.50	788.90	81.62	9.03
SS95 8	134	772.34	743.40	796.60	63.06	7.94
SS96 2	138	675.51	629.80	713.10	219.84	14.83
SS96 3	138	699.24	642.60	729.90	196.50	14.02
SS96 4	138	717.88	666.70	747.00	122.47	11.07
SS96 5	138	731.47	694.20	759.10	109.71	10.47
SS96 6	137	744.91	713.50	779.40	105.30	10.26
SS96 7	135	760.96	730.30	788.60	94.04	9.70
SS96 8	134	774.49	744.70	798.30	73.15	8.55

* * * * * A n a l y's i s o f V a r i a n c e -- design 1 *

* * * *

Orthonormalized Transformation Matrix (Transposed)

	CONST	YR93	YR94	YR95	YR96	GRD3
SS92.2	.169	-.239	.202	-.120	.045	-.254
SS92.3	.169	-.239	.202	-.120	.045	-.169
SS92.4	.169	-.239	.202	-.120	.045	-.085
SS92.5	.169	-.239	.202	-.120	.045	.000
SS92.6	.169	-.239	.202	-.120	.045	.085
SS92.7	.169	-.239	.202	-.120	.045	.169
SS92.8	.169	-.239	.202	-.120	.045	.254
SS93.2	.169	-.120	-.101	.239	-.181	-.254
SS93.3	.169	-.120	-.101	.239	-.181	-.169
SS93.4	.169	-.120	-.101	.239	-.181	-.085
SS93.5	.169	-.120	-.101	.239	-.181	.000
SS93.6	.169	-.120	-.101	.239	-.181	.085
SS93.7	.169	-.120	-.101	.239	-.181	.169
SS93.8	.169	-.120	-.101	.239	-.181	.254
SS94.2	.169	.000	-.202	.000	.271	-.254
SS94.3	.169	.000	-.202	.000	.271	-.169
SS94.4	.169	.000	-.202	.000	.271	-.085
SS94.5	.169	.000	-.202	.000	.271	.000
SS94.6	.169	.000	-.202	.000	.271	.085
SS94.7	.169	.000	-.202	.000	.271	.169
SS94.8	.169	.000	-.202	.000	.271	.254
SS95.2	.169	.120	-.101	-.239	-.181	-.254
SS95.3	.169	.120	-.101	-.239	-.181	-.169
SS95.4	.169	.120	-.101	-.239	-.181	-.085

BEST COPY AVAILABLE

Science scale scores Tennessee 1992 - 1996

SS95.5	.169	.120	-.101	-.239	-.181	.000
SS95.6	.169	.120	-.101	-.239	-.181	.085
SS95.7	.169	.120	-.101	-.239	-.181	.169
SS95.8	.169	.120	-.101	-.239	-.181	.254
SS96.2	.169	.239	.202	.120	.045	-.254
SS96.3	.169	.239	.202	.120	.045	-.169
SS96.4	.169	.239	.202	.120	.045	-.085
SS96.5	.169	.239	.202	.120	.045	.000
SS96.6	.169	.239	.202	.120	.045	.085
SS96.7	.169	.239	.202	.120	.045	.169
SS96.8	.169	.239	.202	.120	.045	.254

	GRD4	GRD5	GRD6	GRD7	GRD8	Y93G3
SS92.2	.244	-.183	.108	-.049	.015	.359
SS92.3	.000	.183	-.252	.195	-.088	.239
SS92.4	-.146	.183	.036	-.244	.221	.120
SS92.5	-.195	.000	.216	.000	-.294	.000
SS92.6	-.146	-.183	.036	.244	.221	-.120
SS92.7	.000	-.183	-.252	-.195	-.088	-.239
SS92.8	.244	.183	.108	.049	.015	-.359
SS93.2	.244	-.183	.108	-.049	.015	.179
SS93.3	.000	.183	-.252	.195	-.088	.120
SS93.4	-.146	.183	.036	-.244	.221	.060
SS93.5	-.195	.000	.216	.000	-.294	.000
SS93.6	-.146	-.183	.036	.244	.221	-.060
SS93.7	.000	-.183	-.252	-.195	-.088	-.120
SS93.8	.244	.183	.108	.049	.015	-.179
SS94.2	.244	-.183	.108	-.049	.015	.000
SS94.3	.000	.183	-.252	.195	-.088	.000
SS94.4	-.146	.183	.036	-.244	.221	.000
SS94.5	-.195	.000	.216	.000	-.294	.000
SS94.6	-.146	-.183	.036	.244	.221	.000

BEST COPY AVAILABLE

Science scale scores Tennessee 1992 - 1996

SS94.7	.000	-.183	-.252	-.195	-.088	.000
SS94.8	.244	.183	.108	.049	.015	.000
SS95.2	.244	-.183	.108	-.049	.015	-.179
SS95.3	.000	.183	-.252	.195	-.088	-.120
SS95.4	-.146	.183	.036	-.244	.221	-.060
SS95.5	-.195	.000	.216	.000	-.294	.000
SS95.6	-.146	-.183	.036	.244	.221	.060
SS95.7	.000	-.183	-.252	-.195	-.088	.120
SS95.8	.244	.183	.108	.049	.015	.179
SS96.2	.244	-.183	.108	-.049	.015	-.359
SS96.3	.000	.183	-.252	.195	-.088	-.239
SS96.4	-.146	.183	.036	-.244	.221	-.120
SS96.5	-.195	.000	.216	.000	-.294	.000
SS96.6	-.146	-.183	.036	.244	.221	.120
SS96.7	.000	-.183	-.252	-.195	-.088	.239
SS96.8	.244	.183	.108	.049	.015	.359

	Y93G4	Y93G5	Y93G6	Y93G7	Y93G8	Y94G3
SS92.2	-.345	.258	-.153	.069	-.021	-.303
SS92.3	.000	-.258	.357	-.276	.125	-.202
SS92.4	.207	-.258	-.051	.345	-.312	-.101
SS92.5	.276	.000	-.306	.000	.416	.000
SS92.6	.207	.258	-.051	-.345	-.312	.101
SS92.7	.000	.258	.357	.276	.125	.202
SS92.8	-.345	-.258	-.153	-.069	-.021	.303
SS93.2	-.173	.129	-.076	.035	-.010	.152
SS93.3	.000	-.129	.178	-.138	.062	.101
SS93.4	.104	-.129	-.025	.173	-.156	.051
SS93.5	.138	.000	-.153	.000	.208	.000
SS93.6	.104	.129	-.025	-.173	-.156	-.051
SS93.7	.000	.129	.178	.138	.062	-.101
SS93.8	-.173	-.129	-.076	-.035	-.010	-.152
SS94.2	.000	.000	.000	.000	.000	.303

BEST COPY AVAILABLE

SS94.3	.000	.000	.000	.000	.000	.202
SS94.4	.000	.000	.000	.000	.000	.101

* * * * * Analysis of Variance -- design 1 * * * *

Orthonormalized Transformation Matrix (Transposed) (Cont.)

	Y93G4	Y93G5	Y93G6	Y93G7	Y93G8	Y94G3
SS94.5	.000	.000	.000	.000	.000	.000
SS94.6	.000	.000	.000	.000	.000	-.101
SS94.7	.000	.000	.000	.000	.000	-.202
SS94.8	.000	.000	.000	.000	.000	-.303
SS95.2	.173	-.129	.076	-.035	.010	.152
SS95.3	.000	.129	-.178	.138	-.062	.101
SS95.4	-.104	.129	.025	-.173	.156	.051
SS95.5	-.138	.000	.153	.000	-.208	.000
SS95.6	-.104	-.129	.025	.173	.156	-.051
SS95.7	.000	-.129	-.178	-.138	-.062	-.101
SS95.8	.173	.129	.076	.035	.010	-.152
SS96.2	.345	-.258	.153	-.069	.021	-.303
SS96.3	.000	.258	-.357	.276	-.125	-.202
SS96.4	-.207	.258	.051	-.345	.312	-.101
SS96.5	-.276	.000	.306	.000	-.416	.000
SS96.6	-.207	-.258	.051	.345	.312	.101
SS96.7	.000	-.258	-.357	-.276	-.125	.202
SS96.8	.345	.258	.153	.069	.021	.303

	Y94G4	Y94G5	Y94G6	Y94G7	Y94G8	Y95G3
SS92.2	.292	-.218	.129	-.058	.018	.179
SS92.3	.000	.218	-.302	.233	-.106	.120
SS92.4	-.175	.218	.043	-.292	.264	.060
SS92.5	-.233	.000	.258	.000	-.352	.000
SS92.6	-.175	-.218	.043	.292	.264	-.060

BEST COPY AVAILABLE

SS92.7	.000	-.218	-.302	-.233	-.106	-.120
SS92.8	.292	.218	.129	.058	.018	-.179
SS93.2	-.146	.109	-.065	.029	-.009	-.359
SS93.3	.000	-.109	.151	-.117	.053	-.239
SS93.4	.087	-.109	-.022	.146	-.132	-.120
SS93.5	.117	.000	-.129	.000	.176	.000
SS93.6	.087	.109	-.022	-.146	-.132	.120
SS93.7	.000	.109	.151	.117	.053	.239
SS93.8	-.146	-.109	-.065	-.029	-.009	.359
SS94.2	-.292	.218	-.129	.058	-.018	.000
SS94.3	.000	-.218	.302	-.233	.106	.000
SS94.4	.175	-.218	-.043	.292	-.264	.000
SS94.5	.233	.000	-.258	.000	.352	.000
SS94.6	.175	.218	-.043	-.292	-.264	.000
SS94.7	.000	.218	.302	.233	.106	.000
SS94.8	-.292	-.218	-.129	-.058	-.018	.000
SS95.2	-.146	.109	-.065	.029	-.009	.359
SS95.3	.000	-.109	.151	-.117	.053	.239
SS95.4	.087	-.109	-.022	.146	-.132	.120
SS95.5	.117	.000	-.129	.000	.176	.000
SS95.6	.087	.109	-.022	-.146	-.132	-.120

BEST COPY AVAILABLE

* * * * * * Analysis of Variance -- design 1 * * * * * *

Orthonormalized Transformation Matrix (Transposed) (Cont.)

	Y94G4	Y94G5	Y94G6	Y94G7	Y94G8	Y95G3
SS95.7	.000	.109	.151	.117	.053	-.239
SS95.8	-.146	-.109	-.065	-.029	-.009	-.359
SS96.2	.292	-.218	.129	-.058	.018	-.179
SS96.3	.000	.218	-.302	.233	-.106	-.120
SS96.4	-.175	.218	.043	-.292	.264	-.060
SS96.5	-.233	.000	.258	.000	-.352	.000
SS96.6	-.175	-.218	.043	.292	.264	.060
SS96.7	.000	-.218	-.302	-.233	-.106	.120
SS96.8	.292	.218	.129	.058	.018	.179

	Y95G4	Y95G5	Y95G6	Y95G7	Y95G8	Y96G3
SS92.2	-.173	.129	-.076	.035	-.010	-.068
SS92.3	.000	-.129	.178	-.138	.062	-.045
SS92.4	.104	-.129	-.025	.173	-.156	-.023
SS92.5	.138	.000	-.153	.000	.208	.000
SS92.6	.104	.129	-.025	-.173	-.156	.023
SS92.7	.000	.129	.178	.138	.062	.045
SS92.8	-.173	-.129	-.076	-.035	-.010	.068
SS93.2	.345	-.258	.153	-.069	.021	.271
SS93.3	.000	.258	-.357	.276	-.125	.181
SS93.4	-.207	.258	.051	-.345	.312	.090
SS93.5	-.276	.000	.306	.000	-.416	.000
SS93.6	-.207	-.258	.051	.345	.312	-.090
SS93.7	.000	-.258	-.357	-.276	-.125	-.181
SS93.8	.345	.258	.153	.069	.021	-.271

BEST COPY AVAILABLE

Science scale scores Tennessee 1992 - 1996

SS94.2	.000	.000	.000	.000	.000	-.407
SS94.3	.000	.000	.000	.000	.000	-.271
SS94.4	.000	.000	.000	.000	.000	-.136
SS94.5	.000	.000	.000	.000	.000	.000
SS94.6	.000	.000	.000	.000	.000	.136
SS94.7	.000	.000	.000	.000	.000	.271
SS94.8	.000	.000	.000	.000	.000	.407
SS95.2	-.345	.258	-.153	.069	-.021	.271
SS95.3	.000	-.258	.357	-.276	.125	.181
SS95.4	.207	-.258	-.051	.345	-.312	.090
SS95.5	.276	.000	-.306	.000	.416	.000
SS95.6	.207	.258	-.051	-.345	-.312	-.090
SS95.7	.000	.258	.357	.276	.125	-.181
SS95.8	-.345	-.258	-.153	-.069	-.021	-.271
SS96.2	.173	-.129	.076	-.035	.010	-.068
SS96.3	.000	.129	-.178	.138	-.062	-.045
SS96.4	-.104	.129	.025	-.173	.156	-.023
SS96.5	-.138	.000	.153	.000	-.208	.000
SS96.6	-.104	-.129	.025	.173	.156	.023
SS96.7	.000	-.129	-.178	-.138	-.062	.045
SS96.8	.173	.129	.076	.035	.010	.068
	Y95G4	Y95G5	Y95G6	Y95G7	Y95G8	Y96G3

SS92.2	.065	-.049	.029	-.013	.004
SS92.3	.000	.049	-.067	.052	-.024
SS92.4	-.039	.049	.010	-.065	.059
SS92.5	-.052	.000	.058	.000	-.079
SS92.6	-.039	-.049	.010	.065	.059
SS92.7	.000	-.049	-.067	-.052	-.024
SS92.8	.065	.049	.029	.013	.004
SS93.2	-.261	.195	-.116	.052	-.016
SS93.3	.000	-.195	.270	-.209	.094
SS93.4	.156	-.195	-.039	.261	-.236

BEST COPY AVAILABLE

SS93.5	.209	.000	-.231	.000	.315
SS93.6	.156	.195	-.039	-.261	-.236
SS93.7	.000	.195	.270	.209	.094
SS93.8	-.261	-.195	-.116	-.052	-.016
SS94.2	.391	-.293	.173	-.078	.024
SS94.3	.000	.293	-.405	.313	-.142
SS94.4	-.235	.293	.058	-.391	.354
SS94.5	-.313	.000	.347	.000	-.472
SS94.6	-.235	-.293	.058	.391	.354
SS94.7	.000	-.293	-.405	-.313	-.142
SS94.8	.391	.293	.173	.078	.024
SS95.2	-.261	.195	-.116	.052	-.016
SS95.3	.000	-.195	.270	-.209	.094
SS95.4	.156	-.195	-.039	.261	-.236
SS95.5	.209	.000	-.231	.000	.315
SS95.6	.156	.195	-.039	-.261	-.236
SS95.7	.000	.195	.270	.209	.094
SS95.8	-.261	-.195	-.116	-.052	-.016
SS96.2	.065	-.049	.029	-.013	.004
SS96.3	.000	.049	-.067	.052	-.024
SS96.4	-.039	.049	.010	-.065	.059
SS96.5	-.052	.000	.058	.000	-.079
SS96.6	-.039	-.049	.010	.065	.059
SS96.7	.000	-.049	-.067	-.052	-.024
SS96.8	.065	.049	.029	.013	.004

Note.. TRANSFORMED variables are in the variates column.
These TRANSFORMED variables correspond to the
Between-subject effects.

BEST COPY AVAILABLE

Table 4

Tests of Year, Grade, and Year by Grade Interaction

		Univariate F	df	Stepdown F	df
YEAR					
	1993	39.84[a]	1	39.84***	1
	1994	17.72[a]	1	0.50	1
	1995	0.65	1	36.34***	1
	1996	1.13	1	32.09	1
GRADE					
	3	11461.53[a]	1	11461.53***	1
	4	443.0287[a]	1	33.91***	1
	5	170.19	1	17.58***	1
	6	1.79	1	10.91***	1
	7	127.79	1	0.41	1
	8	5.28	1	0.84	1
YR by GRD					
	Y93G3	0.09	1	0.09	1
	Y93G4	16.29[a]	1	16.31***	1
	Y93G5	27.04[a]	1	24.37***	1
	Y93G6	12.28[a]	1	9.32**	1
	Y93G7	39.61[a]	1	19.33***	1
	Y93G8	2.09	1	28.47***	1
	Y94G3	2.74	1	0.07	1
	Y94G4	18.98[a]	1	2.66	1

Science scale scores Tennessee 1992 - 1996

		Univariate F	df	Stepdown F	df
	Y94G5	6.73[b]	1	3.43	1
	Y94G6	12.67[a]	1	0.13	1
	Y94G7	22.00[a]	1	12.86***	1
	Y94G8	2.67	1	6.64*	1
	Y95G3	0.07	1	0.11	1
	Y95G4	13.58[a]	1	2.51	1
	Y95G5	17.79[a]	1	1.63	1
	Y95G6	84.33[a]	1	60.57***	1
	Y95G7	48.68[a]	1	25.98***	1
	Y95G8	20.43[a]	1	5.23*	1
	Y96G3	407.09[a]	1	131.20***	1
	Y96G4	23.92[a]	1	2.13	1
	Y96G5	75.70[a]	1	5.32*	1
	Y96G6	0.04	1	0.07	1
	Y96G7	6.52[b]	1	0.36	1
	Y96G8	249.03[a]	1	18.87***	1

Note. [a,b] alpha levels not evaluated.

*p < .051

***p < .001

Table 5

Analysis of Variance Summary Table for Year Effect

Source	SS	df	MS	*F*	η^2
Within	91765.87	528	173.80		
Year	20113.74	4	5028.43	28.93***	0.18

***p < .001

Table 6

Analysis of Variance Summary Table for Grade Effect

Source	SS	df	MS	*F*	η^2
Within	115095.24	792	145.32		
Grade	5063530.13	6	843921.69	5807.24***	0.98

***p < .001

Table 7

Analysis of Variance Summary Table for Year by Grade Effect

Source	SS	df	MS	*F*	η^2
Within	177649.98	3168	56.08		
Year by Grade	54051.70	24	2252.15	40.16***	0.23

***p < .001

List of References

Achilles, C. M., Zaharias, J. B., & Nye, B. A. (1995). Analysis of Policy Application of Experimental Results: Project Challenge . (ERIC Document Reproduction Service No. ED 393 151)

Baker, A. P., & Xu, D. (1995). The Measure of Education: A Review of the Tennessee Value Added Assessment System : Tennessee State Comptroller of the Treasury, Nashville Office of Educational Accountability.

Ballator, N. (1996). The NAEP Guide: A Description of the Content and Methods of the 1994 and 1996 Assessments. Revised Edition: Educational Testing Service, Princeton, N J.

Bock, R. D., Wolfe, R., & Fisher, T. H. (1996). A Review and Analysis of the Tennessee Value-Added Assessment System . Nashville, TN: Office of Education Accountability.

Bruschi, B. A., & Anderson, B. T. (1994). Gender and Ethnic Differences in Science Achievement of Nine-, Thirteen-, and Seventeen-Year-Old Students. (ERIC Document Reproduction Service No. ED 382 751)

Campbell, J. R., & et al. (1996). NAEP 1994 Trends in Academic Progress. Report in Brief: Educational Testing Service, Princeton, NJ Center for the Assessment of Educational Progress ; National Assessment of Educational Progress, Princeton, NJ.

CTB. (undated). Comprehensive test of basic skills spring norms book: March through June . Monterey: CTB.

Finn, J. D., & Achilles, C. M. (1990). Answers and Questions about Class Size: A Statewide Experiment. American Educational Research Journal, 27(3), 557-77.

Hopkins, K. D. (1992). Review of the Comprehensive Test of Basic Skills, Fourth Edition. In J. J. Kramer & J. C. Conoley (Eds.), The eleventh mental measurements yearbook . Lincoln: University of Nebraska, Buros Institute of Mental Measurement.

Johnson, E. G., et al. (1992). The NAEP 1990 Technical Report : National Assessment of Educational Progress, Princeton, NJ.

McLean, J. E., & Lockwood, R. (1996). Why we assess students - and how: The competing measures of student performance. Thousand Oaks, CA: Corwin Press, Inc.

Miller, M. D. (1992). Review of the Comprehensive Test of Basic Skills, Fourth Edition. In J. J. Kramer & J. C. Conoley (Eds.), The eleventh mental measurements yearbook . Lincoln: University of Nebraska, Buros Institute of Mental Measurement.

Miller-Whitehead, M. (1997). An analysis of science scale scores for grades 2-8 in Tennessee for 1990 -1994. Submitted for publication.

Noble, J., & Sawyer, R. (1992). A Comparison of Two Approaches for Measuring Educational Growth from CTBS and P-ACT+ Scores. (ERIC Document Reproduction Service No. ED 346 163)

Norusis, M. J. (1993). SPSS for Windows Base Systems Users Guide Release 6.0. Chicago: SPSS, Inc.

Nye, B. (1993). Some Questions and Answers about Multiage Grouping. (ERIC Document Reproduction Service No. ED 384 998)

Nye, B. A., et al. (1992). Smaller Classes Really Are Better. American School Board Journal, 179(5), 31-33.

O'Sullivan, C. (1995). The Cost of Performance Assessment in Science: The NAEP Perspective: Educational Testing Service, Princeton, N J. (ERIC Document Reproduction Service No. ED 384 638)

Sanders, W. L., & Rivers, J. C. (1996). Cumulative and residual effects of teachers on future student academic achievement (R11-0435-02-001-97). Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Underwood, S., & Lumsden, L. S. (1994). Class Size: ERIC Clearinghouse on Educational Management, Eugene, Oreg ; National Association of Elementary School Principals, Alexandria, VA Office of Educational Research and Improvement (ED), Washington, DC. (ERIC Document Reproduction Service No. ED 377 548)

Yepes Baraya, M. (1995). Task Analysis of Science Performance Tasks and

Items: Identifying Relevant Attributes : Educational Testing Service, Princeton, N J. (ERIC Document Reproduction Service No. ED 388 676)